Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 1 February 2008 (MN LATEX style file v1.4)

Non-Gaussian bounds in the variance from small scale CMB observations

E. Gaztañaga, P. Fosalba, E. Elizalde
Institut d'Estudis Espacials de Catalunya, Research Unit (CSIC), Edf. Nexus-204 - c/ Gran Capitán 2-4, 08034 Barcelona, Spain

1 February 2008

ABSTRACT

We compare the latest results from CMB experiments at scales around $l_e \sim 150$ over different parts of the sky to test the hypothesis that they are drawn from a Gaussian distribution, as is usually assumed. Using both the diagonal and the full covariance χ^2 test, we compare the data with different sets of strategies and find in all cases incompatibility with the Gaussian hypothesis above the one-sigma level. We next show how to include a generic non-Gaussian signal in the data analysis. Results from CMB observations can be made compatible with each other by assuming a non-Gaussian distribution for the signal, with a kurtosis at a level $B_4 = \langle\delta_T^4\rangle_c/\langle\delta_T^2\rangle_c^2 \simeq 90$. A possible interpretation for this result is that the initial fluctuations at the surface of last scattering are strongly non-Gaussian. Another interpretation is that the systematic errors have been understimated in all observations by a factor of two. Other explanations include foreground contamination, non-linear effects or a combination of them.

1 INTRODUCTION

A basic ingredient to understand the formation of large scale structures in our Universe is the distribution of initial conditions. Have fluctuations been generated in the standard inflationary epoch or do they require topological defects or more exotic assumptions for the initial conditions? While the former assumption typically produces a Gaussian distribution (Bardeen, Steinhardt, Turner 1983) the latter involves strong non-Gaussianities (e.g., Vilenkin 1985, Turok & Spergel 1991). This issue can be addressed both in the present day Universe fluctuations, as traced by the galaxy distribution (e.g., Silk & Juszkiewicz 1991, Gaztañaga & Mähönen 1996), or in the the anisotropies of the cosmic microwave background (CMB) (e.g., Coulson et al., 1994, Smoot et al., 1994). Here we will address the latter possibility in a somewhat indirect way. One important contribution to the uncertainties in the measurements of the amplitude of the CMB comes from the sample variance. That is, the uncertainty due to the finite size of the observational sample. In order to estimate these sampling errors it is common practice to assume that the underlying signal is Gaussian (e.g., Bond et al., 1994). These errors are added to other sources of error to test models of structure formation or to compare between experiments. A non-Gaussian signal can produce different sampling errors, and this possibility has already been proposed as a way to reconcile the discrepancies between different experiments (Coulson et al., 1994, Luo 1995)

Here we propose to go a step further and use the estimated discrepancies or variance between different experiments to place bounds on the degree of non-Gaussianity. In order to do this we will assume that the quoted systematic errors in each experiment are accurate, at least on average. We will focus on results which are either from different parts of the sky or, when over the same area, from multipoles with windows that are well separated appart. Our strategy is not to average results from a given experiment, but to find as many independent results as possible in order to have a large sampling over the underlying distribution.

2 SAMPLE VARIANCE

We want to study the sample variance of CMB experiments over independent sky regions or subsamples. We will denote the ensemble average by $\langle\cdots\rangle$. In each subsample we have measurements on several resolution cells or patches, whose averages (within the subsample) we denote by bars. As usual, we assume the *fair sample hypothesis* and in particular that the ensemble averages can be identified with spatial averages (§30 Peebles 1980). In order to derive the sample variance of the temperature fluctuations in the sky for a generic non-Gaussian field, we define our radiation field as $\Delta_m = T_m - \overline{T}$, with T_m the temperature field at a point within certain patch m over which we calculate the subsample average, $\overline{T}$. Notice that the normalized field is given by $(\delta_T)_m = \frac{T_m - \overline{T}}{\overline{T}}$. According to this notation, all magnitudes derived from the field Δ_m may have dimensions. It follows from its definition that the subsample average $\overline{\Delta} = 0$, so that its variance is:

$$\overline{\Delta^2} = \frac{1}{N}\sum_m^N \Delta_m^2 \quad (1)$$

In the literature this quantity is denoted by δT^2, which should not be confused with our notation for the dimensionaless local fluctuation δ_T. The sample variance of $\overline{\Delta^2}$ is therefore the variance of the variance of the temperature field

$$Var\left(\overline{\Delta^2}\right) = \langle\left(\overline{\Delta^2} - \langle\overline{\Delta^2}\rangle\right)^2\rangle = \frac{1}{N^2}\left\{\sum_m \langle\left(\Delta_m^2 - \langle\Delta_m^2\rangle\right)^2\rangle + \sum_{m\neq n}\langle\left(\Delta_m^2 - \langle\Delta_m^2\rangle\right)\left(\Delta_n^2 - \langle\Delta_n^2\rangle\right)\rangle\right\},$$

where we have made use of the fact that the sum commutes with the ensemble average, $\langle\cdots\rangle$. We furthermore assume that the subsamples are large enough to neglect the average cross correlations between patches as compared to the mean square contributions (see Fig.1 in Scott et al 1994). We can then drop the last term and rewrite the first one by commuting back the sum and the averages:

$$Var\left(\overline{\Delta^2}\right) = \frac{1}{N^2}\langle\sum_m\left(\Delta_m^4 - 2\Delta_m^2\langle\Delta^2\rangle + \langle\Delta^2\rangle^2\right)\rangle = \frac{1}{N}\left\{\langle\Delta^4\rangle_c + 2\langle\Delta^2\rangle_c^2\right\}, \quad (2)$$

where we have used that for any X: $\langle\bar{X}\rangle = \langle X\rangle = \langle X_m\rangle$, and we have applied the standard definition for $\langle ...\rangle_c$, connected moments or cumulants (e.g., Kendall, Stuart & Ort 1987).

Throughout the analysis we shall consider a general family of non-Gaussian signals with *dimensional* scaling, which is choosen because it enters at the same level than the Gaussian contribution in the sample variance (see Discussion). For *dimensional* scaling, we have that the 4th order cumulant scales with the square of the 2nd order cumulant, so that:

$$B_4 \equiv \frac{\langle{\delta_T}^4\rangle_c}{\langle{\delta_T}^2\rangle_c^{\ 2}} = \frac{\langle\Delta^4\rangle_c}{\langle\Delta^2\rangle_c^{\ 2}}, \quad (3)$$

is a constant (e.g., independent of $\langle\Delta^2\rangle_c$). In terms of B_4, expression (2) then reads

$$Var(\delta T^2) \equiv Var\left(\overline{\Delta^2}\right) = \frac{1}{N}\,(2 + B_4)\,\langle\Delta^2\rangle_c^2. \quad (4)$$

The Gaussian sample variance corresponds to the particular case $B_4 = 0$. It is important to stress the general applicability of (2) for non-Gaussian processes.

3 SMALL-SCALE CMB DATA COMPILATION

For each CMB experiment over a given subsample, labeled i, we denote as usual $\delta T(i) \equiv \sqrt{\left(\overline{\Delta^2}\right)_i}$ ie, $\delta T(i)$ is the *rms* temperature anisotropy, from which one estimates the band power $\delta T_l(i)$ for every l multipole component of the power spectrum. Table 1 shows a compilation of available data from small-scale experiments for scales within the range $90 \leq l_e \leq 200$. This interval is specially suitable for a χ^2 analysis since it is the most densely sampled, according to observational reports. The scale and size of each window

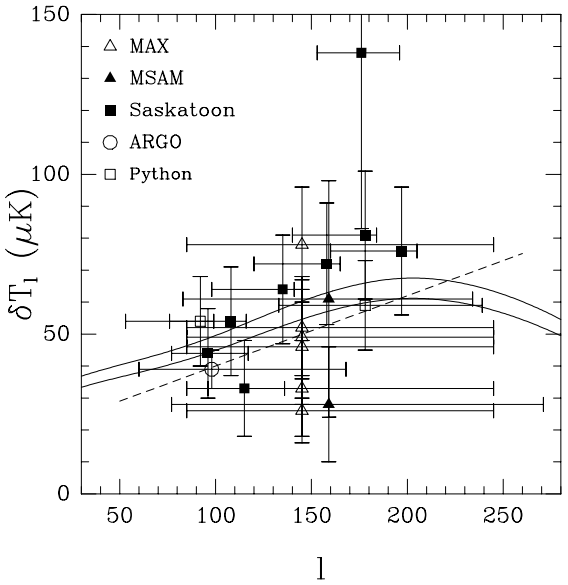

Figure 1. Band power estimates of the *rms* temperature anisotropy δT_l for observations given in Table 1. The vertical error bars show the (symmetrized) total errors in δT_l while the horizontal ones stand for the width of the windows. The dashed line is the best fit slope to the data, $\delta T_l = (11/50)l_e + 18$. Continuous lines show the standard CDM model for two normalizations: $Q_{rms} = 20\mu$ K (top), and 18μ K (bottom).

peaks at multipole number l_e and has a width given by the $\pm\Delta l_e$ interval (computed as the scales at which the window falls to a factor of $e^{-0.5}$ of the peak value). Each input in this table corresponds to independent sky patches or well separated windows. The total quoted error, $\sigma^G_{\delta T}$, includes the calibration uncertainty, the sampling and the instrumental errors. The number of independent points for the statistical analysis is given by the independent bins in RA in each observation. The data esentially follows Ratra et al., 1997, but several cases are taken from the original observational reports. Notice that performing the correct window weighting of the CMB models (e.g., Ratra et al. 1997) hardly changes the final results within the errors, and therefore the flat band hypothesis that we are using for comparing individual experiments should be equally accurate. The data points with their errors (horizontal ones corresponding to the window width) are displayed in Figure 1.

In the above results, Gaussian (G) statistics have been assumed to calculate the sampling variance contribution to the error and this is always included in the quoted errors. For a general non-Gaussian case we would like to replace this contribution with the more general expression given above. The sampling variance estimation in the experiments is usually done with Gaussian Monte-Carlo simulations. Here, to estimate this contribution we will use its theoretical expectation. We first write the rms error σ_{sv} from the sampling variance as: $\sigma_{sv}[\delta T^2] = 2\ \delta T\ \sigma_{sv}[\delta T]$, so that we find from equation (4) that:

$$\sigma_{sv}[\delta T] = \sqrt{\frac{1}{2N}\left(1 + \frac{B_4}{2}\right)}\ \delta T, \quad (5)$$

were N is the number of independent observations. In

Table 1. Available small-scale experimental data within the range $90 \leq l_e \leq 200$. The superscript a denotes the MAX experiments (see Tanaka et al. 1995, and references therein). They are labeled according to the sky patch and flight, MAX-*Sky patch* (*flight*); b denotes the two MSAM1-94 (single difference) experiments reported in Cheng et al. 1996, referring to independent sky regions in RA; c denotes the *Saskatoon*′95 experiments, where SK95Cn correspond to the 95 CAP region, SK94Kn and SK94Qn to the K and Q band experiments in the ′94 flight, with the n point chopping strategy in each case (see Netterfield et al 1997); d, (see de Bernardis et al. 1994); e denotes PYTHON-III$_L$ and PYTHON-III$_S$ for the subtractive large and small chop-window measurements, respectively (see, Platt et al. 1996).

used in Test	CMB experiments	δT_l (μK)	$\sigma^G_{\delta T}$ (μK)	σ^G_{sv} (μK)	l_e	$+\Delta l_e$	$-\Delta l_e$	N_{points}
A1,A2,B	[a]MAX-GUM (3)	78	18	9	145	100	60	39
A1,A2,B	[a]MAX-MUP (3)	26	10	3	145	100	60	39
A1,A2,B	[a]MAX-ID (4)	46	18	7	145	100	60	21
A1,A2,B	[a]MAX-SH (4)	49	19	8	145	100	60	21
A1,A2,B	[a]MAX-HR5127 (5)	33	15	4	145	100	60	29
A1,A2,B	[a]MAX-PH (5)	52	15	7	145	100	60	29
A1,A2,B	[b]MSAM-2 beam(1)	61	37	7	159	75	76	34
A1,A2,B	[b]MSAM-2 beam(2)	28	18	3	159	112	82	34
A2,B	[c]SK95C6	64	17	7	135	6	37	48
A2,B	[c]SK95C7	72	19	7	158	7	38	48
B	[c]SK95C5	54	17	8	108	8	32	24
B	[c]SK95C8	81	20	8	178	6	38	48
B	[c]SK95C9	76	20	8	197	8	37	48
B	[c]SK94K5	44	14	6	96	21	19	24
B	[c]SK94K6	33	15	3	115	21	19	48
B	[c]SK94Q9	138	55	14	176	20	23	48
B	[d]ARGO	39	6	3	98	70	38	63
B	[e]PYTHON-III$_L$	59	14	3	178	61	45	158
B	[e]PYTHON-III$_S$	54	14	3	92	7	39	127

this equation, we have used the individual experiment (or subsample) averages δT instead of the *ensemble* averages: $\langle \Delta^2 \rangle_c^{1/2}$. This is not exact, but reproduces better what is done in each experiment to estimate the Gaussian sampling error, which we denote σ^G_{sv} (i.e., for $B_4 = 0$ above). We have checked that the results shown below do not significantly depend on such approximation. We then assume that the *total* error for a Gaussian signal $\sigma^G_{\delta T}$, given in the observational reports, can be obtained by adding in quadrature σ^G_{sv} to the other errors (e.g., the instrumental and calibration errors). The values of $\sigma^G_{\delta T}$ and σ^G_{sv} for each experiment are shown in Table 1.

We next carry out a Chi-square analysis taking different number of points according to the following:

- **Test A:** Taking a band as narrow and densely sampled as possible, so that we can neglect any dependence of the signal with the scale, l. We consider two cases: **A1, A2**: which correspond to the first 8&10 points of Table 1, respectively.
- **Test B:** Taking a wider band, as densely sampled as possible, and computing the χ^2 value with (**B1**) and without (**B2**) a linear fit to the signal with l_e, i.e. removing a possible scale dependence of the power spectrum in the analysis.

The χ^2 values are to be obtained from the full covariance analysis:

$$\chi^2(c_{ij}) = \sum_{ij} D_i \, C_{ij}^{-1} D_j, \tag{6}$$

were $D_i \equiv \langle \delta T_l \rangle - \delta T_l(i)$ are differences between individual observations (in Table 1) and a theoretical mean value, $\langle \delta T_l \rangle$, which in general varies with l, and $C_{ij} =< D_i D_j >$ is the corresponding covariance matrix. The diagonal terms $C_{ii} = \sigma^2_{\delta T}(i)$, are the individual errors in Table 1. When the D_i's are independent, the covariance matrix becomes diagonal:

$$\chi^2(c_{ii}) = \sum_i D_i^2 \, C_{ii}^{-1} = \sum_i \left(\frac{\langle \delta T_l \rangle - \delta T_l(i)}{\sigma_{\delta T}(i)} \right)^2, \tag{7}$$

The mean $\langle \delta T_l \rangle$ in each test is estimated from the individual values $\delta T_l(i)$ weighted by the inverse of the variance $\sigma^2_{\delta T}(i)$, which produces a minimum χ^2. We have done both a diagonal (7) and a full covariance analysis (6), taking into account the correlations due to calibration uncertainties and the overlap of the window functions. For the off-diagonal terms we use the following:

$$C_{ij} = \kappa_{ij}^2 \, \langle \delta T_l \rangle \, \langle \delta T_{l'} \rangle + w_{ij} \, \sigma_i^{sv} \sigma_j^{sv} \qquad i \neq j \tag{8}$$

The first term correspond to the calibration error, where κ_{ij} is zero for observations i, j in different experiments, and otherwise is the % rms calibration error of the corresponding experiment ($\kappa_{ij} = 0.14, 0.10, .20, 0.10, 0.05$ for Saskatoon, MAX, Phyton, MSAM and ARGO respectively). Note that we need different subscript, l and l', in (8) because the mean can change with l. The second term corresponds to the overlap of the window functions, were σ_i^{sv} are the sampling variance errors (third column of Table 1) and w_{ij} is the normalized overlap between the windows (estimated using $l_e \pm \Delta l_e$) but only when i and j are sampling the same patch of the sky, otherwise is zero. When the mean is definned with the same data, rather than with a theory, the off diagonal terms tend to cancel out, as D_i fluctuates around the mean, but in general the cross-correlations could either increase or decrease the final χ^2.

Table 2 displays the χ^2 values for all the cases involved in Test **A** and **B**. DOF denotes the number of degrees of

Table 2. χ^2 analysis of combined experiments.

Test	Experiments	$\chi^2(w_{ij})$	$\chi^2(\kappa_{ij})$	$\chi^2(c_{ij})$	$\chi^2(c_{ii})$/DOF	$P(\chi^2)$	$\langle\delta T_l\rangle(\mu K)$	B_4
A1	MAX+MSAM				8.4/5	0.14	40.7	$16-238$
A2	MAX+MSAM+SK95C6 & 7				12.2/7	0.09	45.1	$34-276$
B1	ALL	22.5	24.2	23.2	23.9/16	0.09	46.3	$22-140$
B2	ALL+slope	19.2	19.9	19.7	20.0/15	0.17	$0.22l_e+18$	$31-171$
T1	ALL+CDM1	21.9	23.0	22.8	22.4/16	0.12	$C_2=18\mu K$	
T2	ALL+CDM2	29.0	29.7	29.6	29.2/16	0.02	$C_2=20\mu K$	

freedom: $N-3$ (two parameters are correlated to the data: the mean and B_4) for all cases, except for the last case where DOF= $N-4$ —since the slope incorporates one extra parameter to the computation. The values of the χ^2, its probability $P(\chi^2)$ and $\langle\delta T_l\rangle$ shown in Table 2 correspond to the Gaussian case, $B_4=0$, and $\sigma_{\delta T}=\sigma^G_{\delta T}$. In the last test we find that the linear relation that minimizes the χ^2 is: $\langle\delta T_l\rangle = (11/50)\ l_e\ +18$. All cases considered show a disagreement with the Gaussian hypothesis above the 1σ level, e.g., $P(\chi^2)<0.33$, and close to the 2σ level of significance, e.g., $P(\chi^2)\simeq 0.05$.

The values of $\chi^2(c_{ii})$ in Table 2 correspond to the diagonal analysis (7). The full covariance analysis including both terms in (8) is given by $\chi^2(c_{ij})$, while $\chi^2(k_{ij})$ and $\chi^2(w_{ij})$ only takes into account the first or second terms in (8), respectively. The window overlap, although significant in some of the Saskatoon values, hardly makes any difference overall. Thus the full covariance analysis, $\chi_2(c_{ij})$, differs from the diagonal one, $\chi_2(c_{ii})$, by less that 3%, which hardly changes the significance of the analysis below. Therefore, we shall concentrate on the diagonal analysis (7) to take advantage of is simplicity.

For a non-Gaussian signal, the total error above, $\sigma_{\delta T}$, should include the sampling variance for the corresponding non-Gaussian distribution, in equation (5), as well as the instrumental and calibration errors. This can be simply related to the *total* (Gaussian) error $\sigma^G_{\delta T}$, quoted in Table 1, by :

$$\sigma_{\delta T}(i)=\sigma^G_{\delta T}(i)\left\{1+\frac{B_4}{8N_i^2\left[\sigma^G_{\delta T}(i)/\delta T_l(i)\right]^4}\right\}^{1/2}, \qquad (9)$$

which reduces to (5) when there are no systematic errors: $\sigma^G_{\delta T}=\sigma^G_{sv}$. The range for the non-Gaussian parameters B_4 shown in Table 2 are the values needed to produce a χ^2 value corresponding to an interval of confidence between 25%-75%. This range narrows as the number of data points increases, but the mean values are always away from zero.

Note that our approach is not totally consistent. We are assuming a non-Gaussian distribution for the signal but we determine the confidence intervals using the χ^2 distribution, which assumes a Gaussian likelihood. The whole analysis improves substantially by repeating it in terms of a non-Gaussian likelihood function. In the limit $\chi^2\gg N$ and small variance, it is possible to relate the Gaussian confidence intervals with the corresponding non-Gaussian ones in terms of B_4 and B_3^2, where $B_3=\langle\delta_T^3\rangle_c/\langle\delta_T^2\rangle_c^{3/2}$ (see Amendola 1996). Within the limitation that $\chi^2\gg N\gg 1$ (which restricts applicability to Test **B** only), it turns out that the confidence intervals obtained above are widened —when the non-Gaussian corrections are taken into account— by a factor between 1.2 and 2. To do this estimation one has to assume something for the value of B_3^2. The later factors correspond to $B_3=0$ and B_4 in Table 2, which is the most conservative case. This increases the significance to well above 2σ for both of the Test **B** cases in Table 2. Fosalba et al., (1997) have found that typical values of B_3 in several non-Gaussian distributions with *dimensional* scaling lie just below $B_3^2=B_4$. For each assumed value of B_3 we can now find in a consistent way the values of B_4 needed in the χ^2 to get an interval of confidence between 25%-75% in the non-Gaussian likelihood. For $|B_3|=8\lesssim B_4^{1/2}$, the allowed ranges for B_4 corresponding to Test **B1** and **B2** on Table 2 shrink to $B_4=70-90$ and $B_4=90-130$. The improvement in this case is quite remarkable, but the range of allowed values of B_4 increases as B_3 approaches zero.

4 DISCUSSION

Our analysis shows that using the quoted error bars in different CBM experiments, for scales within the range $90\leq l_e\leq 200$, the Gaussian hypothesis can be rejected at a 91% or 83% level, depending on the test (see Table 2). This is not a very significant level, and may also be regarded as 9% or 17% agreement. Our main point is to take these results at face value and see what can be said about non-gaussianities if we seek for a better agreement. We have shown how this can be done for a general class of non-Gaussian distributions by predicting the sampling error. Observations can be made compatible with each other at a 1-sigma level, if the distribution for the signal has a kurtosis at a level of $B_4=\langle\delta_T^4\rangle_c/\langle\delta_T^2\rangle_c^2\simeq 90$. We have repeated the whole analysis taking out each experiment one by one, showing that this result is not dominated by a single measurement. We have also considered subsets of separated experiments (e.g., Test **A1, A2**) and find that this conclusion is robust. As the mean signal is defined from the data to obtain the mimimum χ^2, any comparison with models could only lead to a more significant disagreement. We have also used the code of Seljak & Zaldarriaga (1996) for different CDM universes as the input shape for the mean signal in the χ^2. We normalized the amplitudes according to a quadrupole $C_2=18\mu K$, Test

T1, or $C_2 = 20\mu K$, Test **T2**, as suggested by COBE observations (e.g., Bennet, et al., 1996). These two normalizations of the standard CDM models are shown as the upper and lower continuous lines in Figure 1. The χ^2 and $P(\chi^2)$ values are shown in Table 2. This illustrates that adding curvature to the input model does not reduce the χ^2 values; the linear model is a good approximation to the CDM models for this narrow range of l.

A possible interpretation for this result is that the initial fluctuations at the surface of last scattering are strongly non-Gaussian. Even if this initial distribution were purely Gaussian, it is not clear yet how non-linear effects in the CMB fluctuations or reionization (e.g., Dodelson & Jubas 1995) would change the final observed distributions, although the calculations for some of the relevant effects indicate only mild deviations with cumulants of *hierarchical* type (see e.g., Mollerach et al., 1995, Munshi et al., 1995). Another interpretation is that the systematic errors have been understimated. If we artificially double the systematic errors in *all* experiments at once, we find $\chi^2 = 13$ for Test **B1**, which indicates an agreement at the 67% confidence level. Other possibilities include foreground contamination, which could be in the form of large spots that should typically induce non-Gaussian fluctuations (although de Oliveira-Costa et al., 1997, found this contamination to reduce the Saskatoon normalization by only 2%).

Besides the *dimensional* scaling $\langle {\delta_T}^4 \rangle_c = B_4 \langle {\delta_T}^2 \rangle_c^2$, we have also considered another family of non-Gaussian models: the case of the *hierarchical* scaling mention above, where $\langle {\delta_T}^4 \rangle_c = S_4 \langle {\delta_T}^2 \rangle_c^3$. A similar analysis for the *hierarchical* scaling, yields: $S_4 \simeq 10^{12}$. As the variance δT^2 is of the same order in all data, B_4 just parametrizes $\langle \delta_T^4 \rangle_c$ for any non-Gaussian distribution and the above value agrees well with the naïve expectation: $S_4 \simeq B_4/\delta T^2 \simeq B_4 \times 10^{10}$. Within the large parameter space for non-Gaussian distributions, the values we find for B_4 and S_4 lie in the strongly non-Gaussian cases. In typical non-Gaussian models with *hierarchical* scaling one has $S_4 \lesssim 10^2$ (e.g., Fosalba et al., 1997), much smaller than our result $S_4 \simeq 10^{12}$. For matter fluctuations, δ_m, gravitational growth from Gaussian initial conditions also gives S_4 of order $10 - 100$ (e.g., Bernardeau 1994). For non-Gaussian initial conditions, the topological defects from phase transitions, like textures (e.g., Turok & Spergel 1991), predict $B_4 \simeq 1$, for δ_m, and have been measured around this value in N-body simulations (Gaztañaga & Mähönen 1996), while in the present study we get values around $B_4 \simeq 90$. Thus, the estimated amplitudes for B_4 and S_4 seem to indicate high levels of non-Gaussianity, at least according to δ_m standards. Note that, in principle, one would expect lower levels of non-Gaussianity in δ_T than in δ_m, as the former comes as an integrated effect, at least on large scales (see Scherrer & Schaefer 1995). The large difference in the order of magnitudes between $S_4 \simeq 10^{12}$ and $B_4 \simeq 10^2$ indicates that *dimensional* scaling is a more adequate representation for our findings than the *hierarchical* scaling. If this is the case one would also typically expect a non-vanishing value for B_3 of order $B_3^2 \lesssim B_4$ Thus, we would have $|B_3| \lesssim 10$ or $|S_3| \lesssim 10^6$, much larger than the sampling variance expected in Gaussian models, $\Delta|B_3| \simeq 1$ (e.g., Srednicki 1993). For $B_3^2 \simeq B_4$, we can put a very tight constraint on B_4 to lie between $B_4 = 70 - 90$, for a flat power spectrum, or $B_4 = 90 - 130$ for the best fitted slope in Table 2.

ACKNOWLEDGEMENTS

E.G. and P.F. wish to thank Scott Dodelson for encouraging discussions and the Fermilab group for their kind hospitality during our visit in the fall 1996, were this project started. E.G. acknowledges support from CIRIT (Generalitat de Catalunya) grant 1996BEAI300192. This work has been supported by CSIC, DGICYT (Spain), project PB93-0035, and CIRIT, grant GR94-8001.

5 REFERENCES

Amendola, L., 1996, Astro. Lett. and Comm. 33, 63
Bardeen, J.M., Steinhardt, P.J., Turner, M.S., 1983 Phys.Rev. D,28 679
Bennet, C.L. et al., , 1996, Ap.J.Lett.,464, 1
Bernardeau, F., 1994 A&A,291, 697
Bond, J.R., Crittenden, R., Davis, R.L., Efstathiou, G., Steinhardt, P.J., 1994, Phys.Rev.Lett.,72, 13
Cheng, E.S., et al., 1996, Ap.J.Lett.,456, 71
Coulson, D., Ferreira, P., Graham P., Turok, N., 1994, Nature,368, 27
de Bernardis P., et al., 1994, Ap.J.Lett.,422, 33
de Oliveira-Costa, A., et al., 1997, astro-ph/9705090
Dodelson, S., Jubas, J.M., 1995 ApJ,439, 503
Fosalba, P., Gaztañaga, E., Elizalde, E., 1997, in preparation
Gaztañaga, E., Mähönen, P., 1996, Ap.J.Lett.,462, 1
Kendall, M., Stuart, A. Ort, J.K., 1987, Kendall's Advanced Theory of Statistics, Oxford University Press
Luo, X., 1995, ApJ,439, 517
Netterfield, C.B., Devlin, M.J., Jarosik, N., Page, L., Wollack, E.J., 1997, ApJ,474, 47
Mollerach, S., Gangui, A., Lucchin, F., Matarrese, S., 1995, ApJ,453, 1
Munshi, D., Souradeep, T., Starobinsky, A., 1995, ApJ,454, 556
Platt, S.R., Kovac, J., Dragovan, M., Peterson, J.B., Ruhl, J.E., 1996, astro-ph/9606175
Ratra, B., Sugiyama, N., Banday, A.J., Górski, K.M., 1997, ApJ,481, 22
Scherrer, R.J., Schaefer R.K., 1995, ApJ,446, 44
Scott, D., Srednicki, M., White, M., 1994 Ap.J.Lett.,421, 5
Seljak, U., Zaldarriaga, M., 1996, ApJ,, 469, 437
Silk J., Juszkiewicz, R. 1991, Nature,353, 386
Smoot, G.F., Banday, A.J., Kogut, A., Wright, E.L., Hinshaw, G., Bennett, C.L. 1994, ApJ,437, 1
Srednicki, M., 1993, Ap.J.Lett.,416, 1
Tanaka, S.T. et al., 1995, astro-ph/9512067
Turok N., Spergel, D.N., 1991 Phys.Rev.Lett.,66, 3093
Peebles, P.J.E., 1980, *The Large Scale Structure of the Universe:* Princeton University Press
Vilenkin, A., Phys.Rep.121, 263